

## Durham Research Online

---

### Deposited in DRO:

04 August 2017

### Version of attached file:

Published Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Cliffe, Adam and Doupé, David P. and Sung, HsinHo and Lim, Isaac Kok Hwee and Ong, Kok Haur and Cheng, Li and Yu, Weimiao (2017) 'Quantitative 3D analysis of complex single border cell behaviors in coordinated collective cell migration.', *Nature communications.*, 8 . p. 14905.

### Further information on publisher's website:

<https://doi.org/10.1038/ncomms14905>

### Publisher's copyright statement:

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>. The Author(s) 2017

### Additional information:

## Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

## ARTICLE

Received 4 Nov 2015 | Accepted 10 Feb 2017 | Published 4 Apr 2017

DOI: 10.1038/ncomms14905

OPEN

# Quantitative 3D analysis of complex single border cell behaviors in coordinated collective cell migration

Adam Cliffe<sup>1,\*</sup>, David P. Doupé<sup>1,\*</sup>, HsinHo Sung<sup>1,\*</sup>, Isaac Kok Hwee Lim<sup>1</sup>, Kok Haur Ong<sup>1</sup>, Li Cheng<sup>2</sup>  
& Weimiao Yu<sup>1</sup>

Understanding the mechanisms of collective cell migration is crucial for cancer metastasis, wound healing and many developmental processes. Imaging a migrating cluster *in vivo* is feasible, but the quantification of individual cell behaviours remains challenging. We have developed an image analysis toolkit, CCMTolKit, to quantify the *Drosophila* border cell system. In addition to chaotic motion, previous studies reported that the migrating cells are able to migrate in a highly coordinated pattern. We quantify the rotating and running migration modes in 3D while also observing a range of intermediate behaviours. Running mode is driven by cluster external protrusions. Rotating mode is associated with cluster internal cell extensions that could not be easily characterized. Although the cluster moves slower while rotating, individual cells retain their mobility and are in fact slightly more active than in running mode. We also show that individual cells may exchange positions during migration.

<sup>1</sup>Institute of Molecule and Cell Biology (IMCB), Computational Bioimage Analysis Unit, A\*STAR, Singapore 138673, Singapore. <sup>2</sup>Bioinformatics Institute (BII), Imaging Informatics Division, A\*STAR, Singapore 138671, Singapore. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to W.Y. (email: wmyu@imcb.a-star.edu.sg).

Different types of cells in various contexts migrate together as groups rather than as isolated entities<sup>1,2</sup>. This collective migration of multiple cells is highly directed and coordinated. It is a highly dynamic process involved in immune response, wound healing, tissue development, and cancer metastasis. Many studies of collective cell migration have been undertaken in two-dimensional (2D) tissue culture<sup>3</sup>. Although 2D experiments have provided many insights into general principles, the situation is very different from the endogenous three-dimensional (3D) environment. It has been reported that *in vivo* migration behaviour significantly differs from movement on hard 2D substrates<sup>4,5</sup>. To study cells in a 3D context, we can either make *in vitro* substrates similar to natural conditions or observe collective cell migration directly in the tissue. 3D *in vivo* experiments are the most physiologically relevant but demand the optimization of imaging protocols and advanced image analysis methods.

For *in vivo* studies, 3D time-lapse imaging is becoming less problematic due to advances in fluorescent labelling and microscopy. However, after 3D time-lapses are acquired, a challenging step is to analyse those image stacks using computational approaches to extract meaningful data. The quantitative 3-D analysis should be conducted on relatively large data sets covering multiple movies/cells and extended periods of observation since the biological variation of both the migratory clusters and substrate composition/geometry should be considered. To achieve this requires an automatic, efficient and accurate computational solution to extract relevant quantitative information to better understand the complex behaviours of both the individual cells and the cluster as a whole.

In this paper, we focus on the well-established model of border cells migrating in the *Drosophila* ovary<sup>6</sup>. The migrating cells form a closely packed cluster, comprised of a pair of non-motile hemispheric polar cells, which are surrounded by border cells. The border cells detach from the follicular epithelium and then migrate together between large nurse cells towards the oocyte. The complete process is highly reproducible and takes about 3–4 h. The process is guided by some diffusible cues generated by the oocyte, signaling through the receptor tyrosine kinases PDGF/VEGF receptor (PVR) and epidermal growth factor receptor<sup>7–9</sup>. Previous studies have successfully imaged this process in live isolated egg chambers<sup>7,10</sup>. It requires the homophilic adhesion molecule DE-cadherin on both the border cells and substrate nurse cells<sup>11</sup>. The cytoskeletal regulator RAC also plays a critical role in guiding migration of the border cell cluster<sup>12</sup>. A more recent work suggested a positive mechanical feedback model for such guidance<sup>13</sup>.

In some previous studies, 3D imaging was performed, however, for simplicity, the image data was analysed and interpreted in 2D, either based on single sections or 2D projections of image stacks. Morphodynamics analysis has revealed different patterns of cell extension. However it is still limited to 2D and based on the 2D outer contour of the whole cluster, rather than individual cell in 3D<sup>13</sup>. Overall, this migration process is highly directed, but it contains both chaotic and coordinated movements. Compared to chaotic movement, the coordinated behaviour is more important and relevant to the directed migration. During the coordinated migration, all border cells have similar behaviours, either moving in the same direction (running) or rotating along a given axis (rotating). In general, such coordinated collective movement of cells as a multicellular unit is not fully understood in many biological processes. Furthermore, we do not have the means to quantify these different motions in 3D. At the single cell level, although we know cells display complex behaviours characterized by the cyclic protrusion, adhesion, and contraction of processes in 3D<sup>11–13</sup>, our understanding of the 3D morphology dynamics of every individual cell is limited due to a lack of

suitable computational methods. Here we have studied each cell movement and their morphology within the cluster, providing insights into the mechanisms of coordinated cell movement driving the migration of the cluster as a unit.

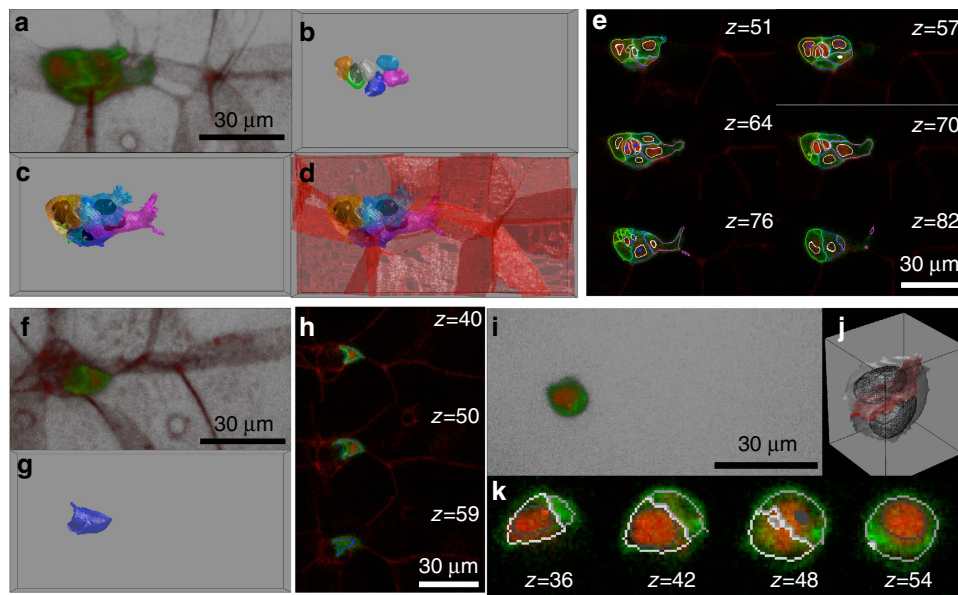
Computational solutions have been reported to address collective cell migration<sup>14–17</sup>, but they are difficult to apply to the border cell system. The challenges of implementing those solutions include the image quality, tracking objects with highly variable speed, and segmentation of the cell surfaces based on weak membrane signal. Overall, a systematic solution designed for this purpose would be extremely useful. For example, a recent development of the ADAPT<sup>18</sup> ImageJ plug-in can analyse cell morphology changes and detect protrusions and blebs, but is again limited to 2D and lacks the capability to separate closely packed cells. In summary, a computational solution to accurately and automatically segment, track and quantitatively describe the complex behaviour of every individual polar/border cell and the cluster in 3D over time is currently lacking.

Here we present an automatic computational method of 3D reconstruction, segmentation, and tracking, enabling the quantitative analysis of migrating border cells on a cell-by-cell basis. Our solution, the Collective Cell Migration Toolkit (CCMToolKit), is freely available at <https://sites.google.com/site/ccmtoolkit/>. Quantitative information, such as speed, volume, morphology and movement trajectories of each nucleus and cell, are extracted. The obtained per-cell based information provides us with an extensive detail on how the cells are coordinated together and move forward. In addition to the conventional parameters for individual cells' behaviour, we also quantify cell–cell interactions within the cluster, for example, the interacting surface between two border cells and the degree of coordination. We also quantify neighbour exchange as adjacent border cells change their relative positions with each other within the moving cluster.

This numerical solution enables us to address some open questions, many of which are difficult or impossible to characterize in 2D. To quantify different types of coordinated cluster movement, we provide two parameters, Group Polarization (GP) and Angular Momentum (AM), to classify directed and tumbling motions of the cluster, that is, the running and rotating modes, which were qualitatively observed<sup>2,5,6,9,11–13</sup>. A numerical classification boundary is provided to differentiate these two modes, but there is a continuum of behaviours between them. We also show that the polar cells usually rotate together with the cluster, but individual border cells can slide along the surface of polar cells in a process we refer to as neighbour exchange. We provide a numerical metric (temporal topology change metric) to detect this neighbour exchange automatically. Clusters generate protrusions during both running and rotating modes, but it has been difficult to quantify their frequency and contribution to the two modes. A computational method is presented to detect cluster protrusions and quantify their role in both running and rotating. Assuming the temporal interval of image acquisition is sufficiently small, we may consider the deformation of the cell surface between two time points as linear. We therefore propose a method to calculate the deformation of each cell according to linear points mapping, offering new insights into cell extension within the cluster and its association with cluster rotation and neighbour exchange. Finally, we address the question of whether individual cells move more slowly during rotating mode.

## Results

**Segmentation and reconstruction of migrating cell cluster.** 3D time-lapse movies of migrating border cell clusters were generated as described in Methods. We use pre-processing steps



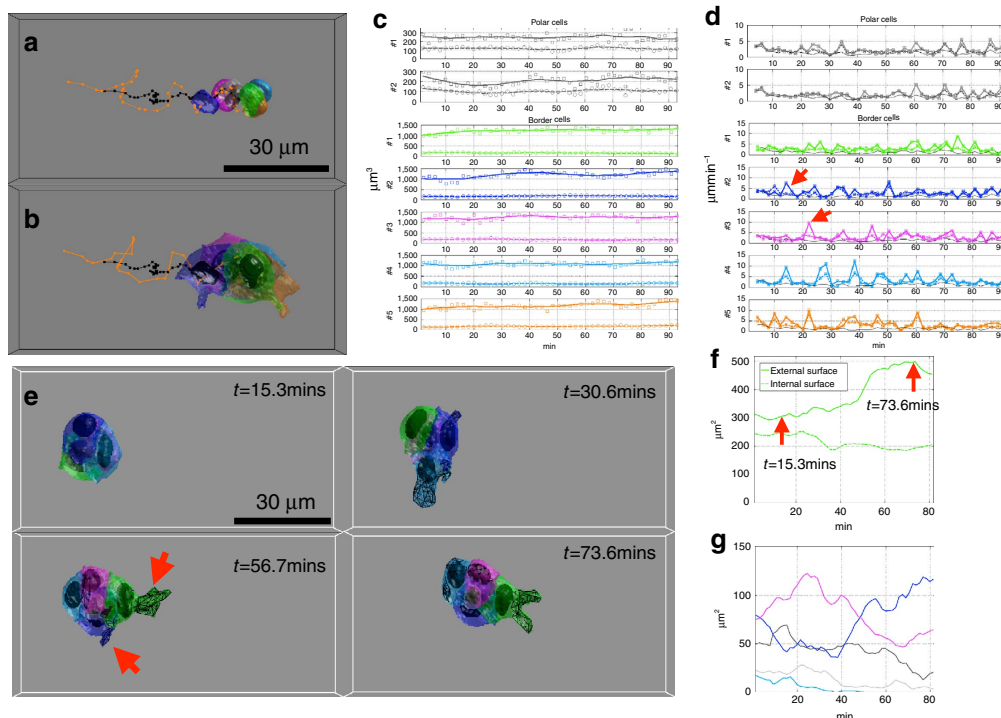
**Figure 1 | Segmentation and reconstruction of migratory cells and nuclei in the cluster.** (a) Z-projection of an original image stack at a single time point showing migrating cell nuclei (red), membranes (green) and nurse cell membranes (red). (b) Segmented nuclei with different cell identities in different colors. Two polar cells are in light and dark grey. (c) Reconstructed migrating cell surfaces in 3D with nuclei in black and surface colors as in b. (d) Reconstructed migrating cells/nuclei and nurse cell surfaces (red). (e) Cell boundaries of colors as in c and nuclei outlines (white) at different z-slices of a cluster in a. (f) Original image stack of single border cell labelling with border cell membrane (green) and nurse cell membranes (red). (g) Reconstructed single border cell surface in blue. (h) Slice view of the original signal overlaid with the border cell contour at different z-slices. (i) Original image stack of polar cell labelling. (j) Reconstructed polar nuclei surfaces (meshed surface) and cell surfaces (light and dark grey). Red indicates the interacting surface between two polar cells. (k) Slice view of the original signal with the polar cell surfaces (light and dark grey) overlaid at different z-slices.

to address the problems of photo-bleaching, anisotropic voxel size, poor signal-noise-ratio and the mixture of signals (the fluorescent protein used to label the border cell nuclei has very similar emission spectrum to the nurse cell membrane marker), see Methods section for more details. The image stacks are reconstructed and enhanced at each time point. The photo-bleaching is corrected as shown in Supplementary Fig. 1. We segment the migrating cell nuclei based on the Gaussian Mixture Model (GMM). A dedicated rotational kernel follows this to enhance the weak membrane signal further. To reduce the computational cost, we rotated a kernel in  $x$ - $y$ ,  $x$ - $z$ ,  $y$ - $z$  planes, convoluted the kernel with membrane signal and enhanced the membrane signal based on the maximum response of the filter bank. A similar membrane enhancement scheme was reported by Mosaliganti *et al.*<sup>19</sup>. A marker-controlled watershed approach is then applied to segment the cell membranes.

Our method reconstructs the collective migration of border cells in 4D (3D and time), for both cells and nuclei, with high accuracy and reliability. We resolve the surface of each nucleus, cell and its nurse cell substrates as shown in Fig. 1a–d. We overlaid the cell surface (colour-coded contour) with the original image to visually inspect the segmentation quality, and representative z-slices are shown in Fig. 1e. In our segmentation, it is crucial to constrain the cluster volume as a constant, such that the volume of each cell is optimized. To define border/polar cell volume, we acquired movies where only one cell within the cluster is labelled as shown in Fig. 1f,g. The boundaries of a border cell are overlaid with the original image in Fig. 1h. Example data of polar cells are shown in Fig. 1i–k with the interacting surface between the two polar cells visualized in Fig. 1j. These specifically labelled image stacks provide us with cell volume and based on that we constrain the total cluster volume according to the cell number. The boundaries between the cells are determined by the enhanced membrane signal and

the marker-controlled watershed algorithm. Accurate segmentation facilitates the nucleus/cell tracking that follows and is a critical step to investigate the behaviour of each cell within the migrating cluster. The 4D reconstruction of a moving cluster, including both nuclei and cell surfaces, are shown in Supplementary Movie 1a–b and a segmentation result at different z-slices is shown in Supplementary Movie 2. For the purpose of validating our computation results, 10 cells (with ~10–39 time points for each cell) were randomly selected from different movies and manual segmentation was performed. Our computational results are highly consistent with manual segmentation as illustrated in Supplementary Fig. 2. Overall, the true positive rate is >82%, the false negative is about 15% and the false positive is about 30%. A visual comparison shows that false positives are mainly due to manual segmentation failing to capture fine cell structures. Notably, our computational results show much smaller cell volume fluctuations over time than the manual methods in Supplementary Fig. 3 indicating that computational segmentation is more consistent than manual segmentation.

**Tracking nuclei/cells and cell-cell coupling.** The defined tracking cost function in Methods guarantees a global minimum and is able to handle relatively large displacements of nuclei/cells during migration. The tracking analysis provides us with the dynamics of each migrating cell and its nucleus in 4D. The mass center trajectories of one cell and its nucleus are illustrated in Fig. 2a,b, respectively. The cluster mass center trajectory is also shown in black. It is clear that although the cluster movement is relatively straight, an individual cell/nucleus does not move in strict straight line. The volumes of each polar and border nucleus/cell are presented in Fig. 2c. In this panel and Fig. 2d (results of nuclei/cell speed), grey colour indicates two polar cells



**Figure 2 | Quantitative tracking of migrating cells and cell-cell coupling.** (a,b) Tracking of migratory nuclei (a) and cells' (b) mass center. Different colors indicate cell identities. The trajectories in orange are the moving path of the orange cell (b) and its nucleus (a). Although cluster mass center moves relatively straight and directed (as shown by the black trajectory), each cell does not move as straight as the cluster center and has individual behaviour within the cluster. (c) Quantification of nucleus (dashed line) and cell (solid line) volume during migration for each cell. The color code indicates the cell identities as in a,b. The nuclei and cell volumes are relatively stable. (d) Velocity magnitudes of each individual nucleus (dashed line) and cell (solid line). Cells are mostly at low speed ( $<5 \mu\text{m min}^{-1}$ ), but at some time points can be at much higher speed ( $\sim 10 \mu\text{m min}^{-1}$ ). (e) Morphology of the cells and cluster at four different time points. (f) Dynamics of internal and external surfaces of the green cell in e. Although the internal surface is almost a constant, the external surface increased about 40% at the latter stage of this movie highlighted by the red arrows. (g) Cell-cell interacting interface between the green cell and other color-coded cells as in e. The green cell initially has more interaction with the magenta cell, while the interacting interface with the blue cell approximately doubles in the latter half of the movie.

and other colours are individual border cells as in Fig. 2a,b. As we already knew there is no mass exchange between cells during the migration and border cells do not divide, the volume of each cell is relatively stable (Fig. 2c). By analysing the speed of both nuclei and cells (Fig. 2d), the border cells usually move at low speed ( $<5 \mu\text{m min}^{-1}$ ), but display short bursts of higher speed  $\sim 10 \mu\text{m min}^{-1}$ . This indicates migrating border cells can frequently exhibit individual, uncoordinated intra-cluster motility. The speeds of nuclei are slightly lower than the cells and show less fluctuation (Fig. 2d and Supplementary Fig. 4) but they are highly correlated (Supplementary Fig. 5). In general, polar cell and nucleus speed is lower than that of border cells/nuclei, consistent with a more passive role in migration.

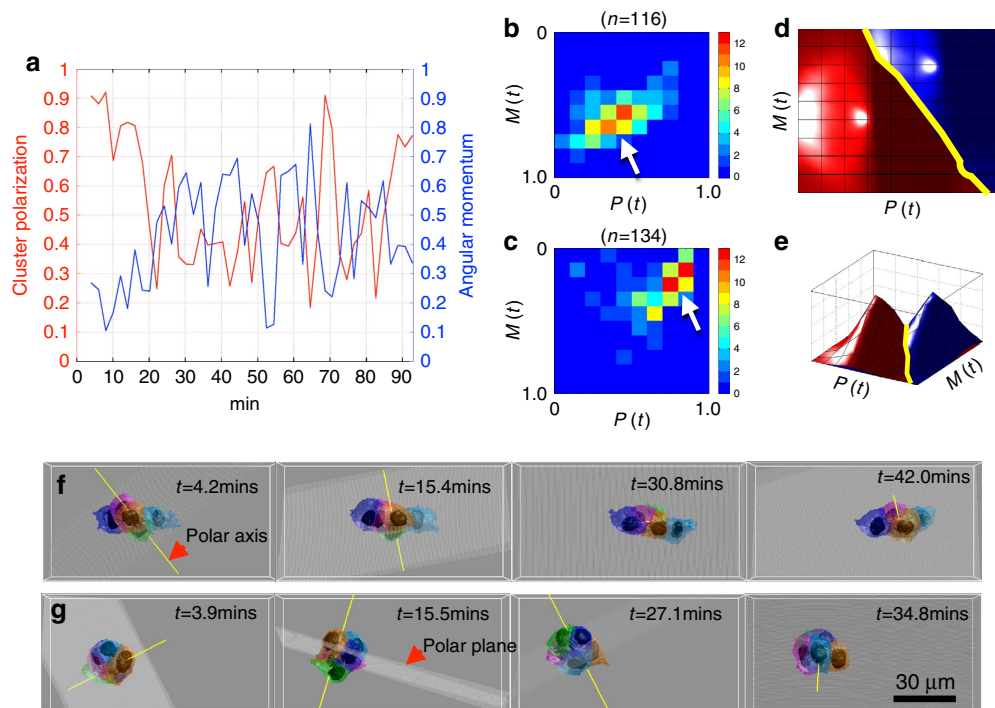
In previous works, the border cell velocity was measured using 2D projections at  $\sim 1.5 \mu\text{m min}^{-1}$ . Here segmentation and tracking individual cells accurately in 3D, allows us to quantify border cell speed in 3D space. The average speed of polar and border cell is about  $2\text{--}3 \mu\text{m min}^{-1}$ . Our reported value is higher than previous 2D reports because the speed along the z-axis is now taken into account. Indeed, when we remove speed in the z-direction our results show an average speed of  $1.7 \mu\text{m min}^{-1}$ , consistent with previous reports. Although the movement of border cells along the z axis does not contribute directly to the cluster reaching its destination, it allows us to better understand the overall border cell behaviour in 3D. As polar cells form the center of the cluster and do not touch the nurse cells, their

movement is analogous to the net cluster movement (Supplementary Fig. 6).

The reconstruction of each cell's surface in 3D at different time points allows us to look into how the cells are interacting with each other to achieve coordinated movement. Each cell typically has two types of interface: external interface, the surface touching the substrate, that is, the nurse cells; and internal interface, the surface interfacing with other migrating border cells or polar cells.

To calculate the external and internal surface of one given cell, we dilate it by a  $3 \times 3 \times 3$  structural element, effectively making that particular cell slightly 'fatter'. On the basis of the areas that now overlap with other cells, we may then estimate the interacting surface. Although the volume of a border cell is relatively stable, its external and internal interfaces may vary significantly. In Fig. 2f, we show that the green cell in the Fig. 2e cluster has less interaction with nurse cells at the beginning, that is, a smaller external interface, but this contact area increases about 40% at the latter stages of this movie while the internal interface remains approximately constant. The increase in the external interface is closely associated with the generation of a large cluster external protrusion in the second half of this movie. The interface between border cells, which was not accessible, is a key parameter to understand the cell-cell interaction and coordination during the collective movement. The area and variation of these interfaces may indicate the strength of binding between cells and the overall interacting





**Figure 3 | Coordination of migratory cells: running and rotating modes.** (a) The measurements of group polarization and angular momentum of a movie over time. The values of group polarization and angular momentum are within [0, 1]. We can project to a 2D space and classify the cluster movement into running and rotating. (b,c) Probability distribution of group polarization and angular momentum of four movies with purely rotating (b) or running (c). There are 116 time points for cluster rotating and 134 time points for running. In this 2D space, we can clearly see two distinct motion regimes, indicated by white arrows. (d,e) GMM fitting of group polarization and angular momentum of rotating (red surface) and running (blue surface). The decision boundary of these two modes, indicated by the yellow line, is the edge between the red and blue surface. (f) Polar axis (yellow) and plane (light grey) for a running movie. The relative position of the two polar cells is relatively stable during the running mode (Supplementary Movie 5). (g) Polar axis and polar plane for a rotating movies. The polar cell rotates with the whole cluster during the rotating mode. (Supplementary Movie 5).

patterns between all cells define the coordination of the cluster. Such analysis allows the study of the interactions between border cells and shows that the cluster internal interaction between cells can change dramatically during the migration (Fig. 2g).

### Running versus rotating modes and polar cell orientation.

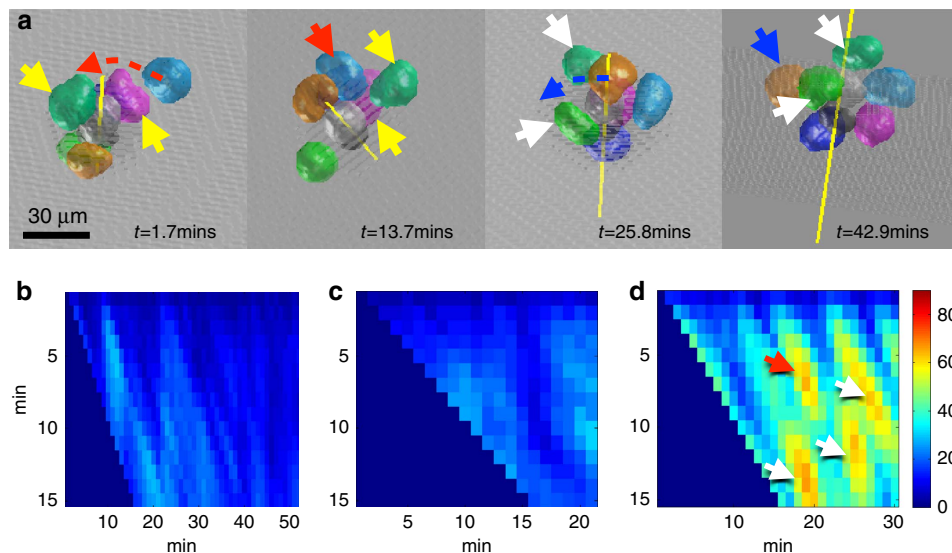
Although the cluster exhibits random and chaotic movements, highly coordinated modes of movement, running and rotating, have been qualitatively observed<sup>2,5,6,9,11–13</sup>. However, the mechanism, which governs the coordination between the cells, is not clear and we also lack quantitative parameters to quantify and classify these two different modes. We define two features, group polarization (GP) denoted by  $P(t)$  and angular momentum (AM) denoted by  $M(t)$ , to quantify the coordination. Full details of these two parameters are provided in Methods. Briefly  $P(t)$  describes how well the movement of each nucleus/cell is aligned within the cluster. When all cells move in a similar direction, it indicates a high value of  $P(t)$ , as can be seen in running mode. Similarly,  $M(t)$  tells us if the cells are rotating in a similar direction. The values of both  $P(t)$  and  $M(t)$  of a movie are illustrated in Fig. 3a.

During certain periods of cluster migration, border cells exhibit highly coordinated movements, as either all cells are moving towards their destination or rotating around the polar cells. We can quantify them as two distinct migration modes based on  $P(t)$  and  $M(t)$ . In running mode,  $P(t)$  has higher value, but  $M(t)$  has smaller value; whereas in rotation mode  $P(t)$  has smaller value, but  $M(t)$  has higher value. In Supplementary Movie 3–4, we demonstrate the difference between the running

and rotating modes. In these videos, the yellow line is so-called polar axis, which is a straight line passing through the geometric centers of the two polar cells. This allows us to visualize if the polar cell is rotating with the border cell when a cluster is rotating.

We selected 4 movies, which apparently exhibit either only running (134 time points) or rotating (116 time points) modes. We calculate  $P(t)$  and  $M(t)$  and then project their distribution in  $P(t)$  versus  $M(t)$  space for rotating movies (Fig. 3b) and running movies (Fig. 3c). The local maxima in Fig. 3b is shifted in Fig. 3c, indicated by the white arrows. The cluster forward speed difference between running and rotating is shown in Supplementary Fig. 7. Although rotating mode produces little cluster forward speed, cells are still well coordinated within the cluster.

Although we may quantify the coordinated movement as running and rotating modes, we need a general principle to differentiate these two modes and decide whether a cluster at a given time point is under running mode or rotating mode based on these two parameters. We fit our data using a 2D Gaussian model in Fig. 3d,e, where red surface indicates rotating and blue surface indicates running. The yellow boundary in Fig. 3d defines its separating surface between running and rotating modes. Supplementary Fig. 8 shows the distribution of the two features for 42 migrating clusters. The classification of running and rotating mode can be achieved by applying the decision boundary indicated by the yellow line in Fig. 3d and Supplementary Fig. 8. In Supplementary Fig. 8, most of the cluster movements are located at the diagonal of the matrix of  $P(t)$  and  $M(t)$ , indicating that overall the cluster



**Figure 4 | Neighbour exchange and its association of cluster internal cell extensions.** (a) Two incidences of neighbourhood exchange in a movie. The light blue cell (represented here by its nucleus), whose movement is indicated by the red dotted arrow at  $t=1.7$  m, penetrates between the green and magenta cells, indicated by the yellow arrows at  $t=1.7$  m and  $t=13.7$  m, within  $\sim 10$  m. The gold cell, whose path is indicated by the blue dotted arrow, has similar behaviour and moves between the two green cells indicated by the white arrows. (Supplementary Movie 6). (b,c) The temporal topology change metric of two clusters, running and rotating respectively, without neighbour exchange. The reconstructed movies are shown in Supplementary Movie 7 and 8, respectively. (d) The temporal topology change metric of a cluster with four neighbour exchanges. The reconstructed movies are shown in Supplementary Movie 6. The x axis of hot peaks indicates the time points of such incidences occurring. The y axis indicates how long this neighbour exchange will take in min.

are highly coordinated. However, it is clear that there is a continuum of behaviours and some will not fall neatly into either category, that is, clusters with low  $P(t)$  and  $M(t)$ . In Fig. 2d, we show that the border cells move at low speed ( $< 5 \mu\text{m min}^{-1}$ ) most of the time, but display short bursts of higher speed  $\sim 10 \mu\text{m min}^{-1}$ . These bursts are often not seen in all cells/nuclei in the cluster at a given time (see red arrows in Fig. 2d). This shows that although cells usually move in a highly coordinated fashion as in Supplementary Fig. 8, border cells can still frequently exhibit individual, uncoordinated intra-cluster motility. These intermediate states and the transitions between chaotic and coordinated movement will be of interest for further investigation.

Another open question we can interrogate is whether the polar cells are rotating together with the border cells or remaining in the same orientation while border cells slide over the polar cell surfaces. In Fig. 3f,g, we visualize the polar axis and the polar plane, a plane perpendicular to this axis. We show that the polar axis generally rotates together with the whole cluster during rotating mode in Fig. 3g but is very stable during the running mode. Supplementary Movie 5 clearly demonstrates the difference in polar axis rotation between the running and rotating modes. We have quantified this by taking the moving average of polar axis rotating angles (over 5 time-points) for both running and rotating movies (Supplementary Fig. 9).

The cell–cell interaction surface is an indicator of how the cells are coordinated, and we wondered if it is different in running and rotating modes. Supplementary Fig. 10 shows that the interaction surface is more consistent during running (Supplementary Fig. 10a) than rotating (Supplementary Fig. 10b).

**Neighbour exchange during the cluster migration.** Both the whole cluster behaviours and the individual cell behaviours have been observed to be highly dynamic. A cell may occupy a leading

position at one time point, then be found at the rear or side of the cluster later. This process could occur by rotation of the whole cluster as already reported in refs 2,13. Another scenario is that a cell actively exchanges position with its neighbours, which we term neighbour exchange. Two such examples are shown in Fig. 4a. The light blue cell (represented by its nucleus), moving along the path of the red dotted arrow at  $t=1.7$  m, passes between the green and magenta cells indicated by the yellow arrows, and reaches the rear of the cluster within  $\sim 10$  min. The gold cell, moving along the blue dotted arrow, has similar behaviour and moves between the two green cells indicated by the white arrows. Supplementary Movie 6 illustrates this example of the neighbour exchange. At time point = 23 in this video, the shape of the gold cell nucleus is elongated probably because it was squeezed by the two green cells. During the neighbour exchange, the topological arrangement of the other cells remains largely unchanged. We can glean some information about the relative position of a cell by observing the interface with its neighbours. However, this is insufficient to fully describe neighbour exchange. To obtain a complete picture of moving cluster, we must identify neighbour exchange based on global topological change. We designed a temporal topology change metric in Methods and results are shown in Fig. 4b–d. The x axis shows when a neighbour exchange happens and the y axis shows how long it will take. Results for two representative rotating and running movies without any neighbour exchange are shown in Fig. 4b,c (Corresponding reconstructed movies are shown in Supplementary Movie 7 and 8, respectively). A movie with neighbour exchange is shown in Fig. 4d (Reconstructed movie shown in Supplementary Movie 6). The local maxima in Fig. 4d, the hot spots indicated by the annotated arrows, show that this movie has four events of neighbour exchange. The local maxima highlighted by the red arrow in Fig. 4d suggests that this neighbour exchange takes about 6–7 time points at around time-point = 20 (The gold cell nuclei in Fig. 4a). Overall, the relative positions of the border-

polar cells are stable during both running and rotating, but it is possible for border cells to slide along the surface of polar cells and exchange their position with other border cells.

**Cluster external protrusion and internal cell extension.** During cluster migration, border cells may generate elongated protrusions extending out along the norm direction from the main body of cluster (Fig. 2e meshed surface highlight by the red arrows). The definitions of different types of cell protrusions are summarized in ref. 4. However, there are not yet clear mathematical or morphological criteria to differentiate them. We describe a morphology operation in Methods to detect and measure these functional structures. Our computationally defined protrusions may not exactly correspond to cell biology defined F-actin filled protrusions. We designate them as cluster external protrusions but without further classification into different sub-categories. The cluster external protrusion is a dynamic structure at the whole cluster level and it can take up to 50–70% volume of a border cell, potentially including the volume of the nucleus.

On the basis of our segmentation and tracking, it is possible to accurately describe the deformation of an individual border cell over time. The dynamics of cell surface deformation was previously inaccessible but is a key aspect to understand single cell behaviours in a cluster. To quantify this deformation over time, we assume the sampling interval of image acquisition is small enough and thus provide a linear mapping solution in the Computational 3D Morphological Dynamics section of Methods. Besides the massive cluster extensions, border cells display more individual behaviours and can generate cell extensions along the tangent direction of the cluster. We summarize these two types of protrusions as follows. A cluster external protrusion is an elongated protrusion detected using morphology operation along the norm direction of the cluster. It often makes contacts with two or more nurse cells but not other border cells. Typically, these protrusions are visible at the cluster level, squeezing in between the interface of nurse cells. A cluster internal cell extension is an individual cell extension, detected by computational deformation analysis. Individual cells produce extensions along the tangent direction of the cluster. These extensions normally occur at the interface between other border cells and nurse cells. In this case, a border cell is observed ‘reaching around’ another cell. Such behaviour plays an important role in the neighbour exchange events as illustrated by Supplementary Movie 9.

**Cluster external protrusion and running mode.** Cluster external protrusions are illustrated by the meshed surface in Fig. 5a and Supplementary Movie 9. The morphology dynamics of the cluster external protrusion is shown in Fig. 5b and Supplementary Movie 10. Using our method, we can identify which cell produced the protrusion and quantify the protrusion volume ( $\mu\text{m}^3$ ) and its direction. The generation of cluster external protrusions is very different during running and rotating modes in Fig. 5c. The average cluster external protrusion rate is  $\sim 25 \mu\text{m}^3 \text{min}^{-1}$  for rotating mode, while the cluster can generate massive cluster external protrusion at  $150 \mu\text{m}^3 \text{min}^{-1}$  during running mode.

In what follows we provide analysis on the different roles of cluster external protrusion in running and rotating modes as presented in Fig. 5d. The first parameter is the alignment of each cluster external protrusion direction and the cluster migrating direction. The protrusion directions are first normalized, that is, only considering the direction as a unit vector without magnitude information. Each vector conducts a dot product with the cluster direction (also normalized). We take the magnitude of the dot product as protrusion alignment (PA) with cluster direction in Fig. 5d. If the protrusion direction perfectly aligned with cluster

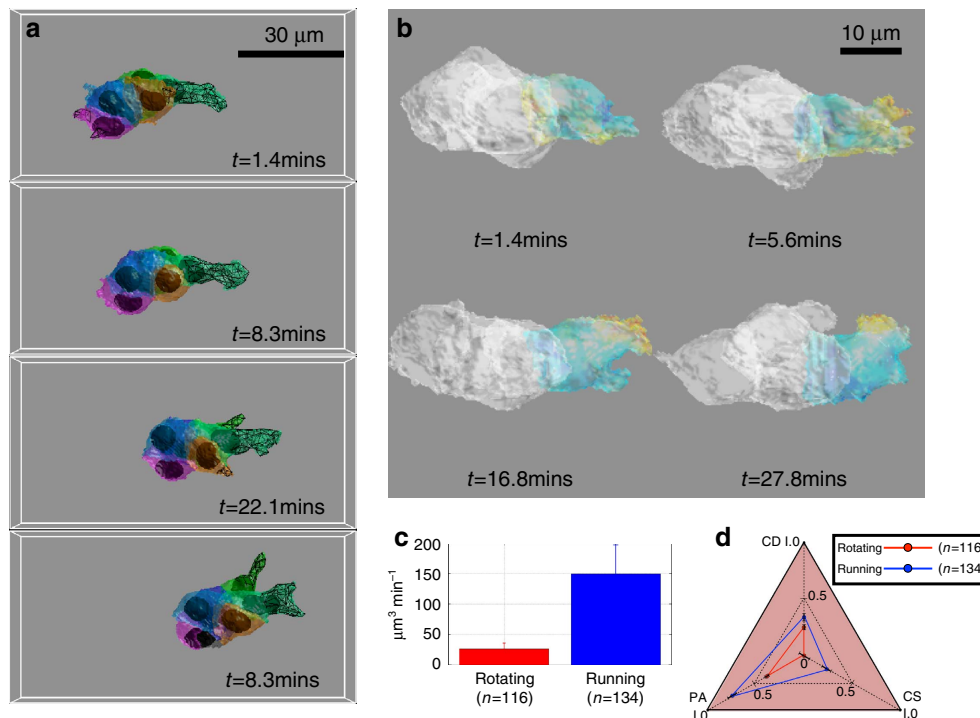
migration direction,  $\text{PA} = 1$ . When opposite,  $\text{PA} = -1$ . During rotating mode, the cluster external protrusion direction can show correlation with the cluster movement direction, that is, the PA is  $< 0.4$ , the correlation is much higher during running mode, PA is  $> 0.7$ . The normalized protrusion direction vectors are added as a combined vector. Then we applied correlation analysis between this vector’s magnitude and the cluster speed magnitude as the impact to cluster direction (CD) in Fig. 5d, that is, combined directions versus cluster speed magnitude. Then we conducted correlation analysis between the two variables (see Computational 3D morphological dynamics in Methods section). This parameter indicates whether the combined protrusion direction correlates with the cluster speed magnitude. CD is also higher during running mode in Fig. 5d. External protrusions are therefore less related to forward movement during rotating than running mode. Finally, we measure the correlation between the protrusion mass and the cluster speed magnitude, as the impact of protrusion size to cluster speed (CS). In rotating, the cluster external protrusion has almost no correlation to cluster movement, while it much more highly correlated during running mode as shown by CS in Fig. 5d.

**Cluster internal cell extensions and rotating mode.** The deformation of individual cells within the cluster may provide us with additional insights into cluster coordination. In the Computational 3D Morphological Dynamics section of the Methods, we provided a linear mapping solution to estimate the deformation of a given cell between two consecutive time points. A 2D illustration of our linear mapping is given in Fig. 6a. The actual 3D deformation of a border cell is illustrated by color-coded vectors in Fig. 6b. Hot color indicates positive deformation and cold color indicates negative deformation. During the running mode, the leading cell front edge is also very active, as shown in Supplementary Movie 11. Similarly, in the rotating mode, the cluster internal protrusion is detected by the positive deformation as indicated by the blue arrows in Fig. 6c. During rotating mode, the cells generate cell extensions between the interface of the nurse cell and border cells as seen in Supplementary Movie 12. The morphodynamics and the deformation measurement of a border cell within a rotating cluster are shown in Fig. 6c. The cluster internal cell extensions are associated with the cluster rotating direction indicated by the red dotted arrows. We hypothesize that uncoordinated cluster internal extensions, as illustrated in Supplementary Movie 7, would also provide a mechanism for individual and uncoordinated intra-cluster motility, such as the described neighbour exchange.

Supplementary Fig. 11a,b show the positive deformation of all cells in the running and rotating cluster respectively. The positive deformations are shown in color-coded solid line and the negative deformations are in the dashed color line. The deformation of the front (positive deformation) is much bigger than rear edge (negative deformation) of a given cell in both running and rotating modes. We also observed that the dynamics calculation, positive deformation energy (PDE), as defined in Methods, and shown in Supplementary Fig. 11, has a periodic behaviour with a frequency of  $\sim 3$ –4 min. However, our current temporal resolution does not allow a conclusive statement. Better temporal and spatial resolutions are required for further investigation.

Consistent with previous reports we show that the cluster speed is lower in rotating mode than running (Supplementary Fig. 7). However, it is not clear if the individual cells move more slowly during the rotating mode. Our computational analysis reveals, counter-intuitively, that the cells are slightly more active in the rotating mode than the running mode based on 6 parameters defined in Methods section (Fig. 6d).





**Figure 5 | Cluster external protrusions are associated with cluster running.** (a) Cluster external protrusion detection and measurement at four different time points. The meshed surface indicates the cluster external protrusion and the color indicates which cell generates the cluster external protrusion. (b) The morphological dynamics of a leading cell of a running cluster ( Supplementary Movie 11). Hotter color indicates bigger deformation. (c) Quantification of cluster external protrusion in rotating and running modes. During rotating mode, the cluster generates  $\sim 25\ \mu\text{m}^3\ \text{min}^{-1}$  protrusions, while during running mode it can generate much larger protrusions of  $150\ \mu\text{m}^3\ \text{min}^{-1}$ . Error bar is the Standard Error of the mean. (d) Cluster external protrusion parameters differ during rotating and running modes. PA: alignment of each cluster external protrusion with cluster migrating direction, protrusion directions are first normalized, and the magnitude of the unit vector's dot product with the cluster direction (also normalized) is taken as PA. In running phase, the protrusions are more highly correlated with the cluster direction, i.e. the cluster moves along the cluster external protrusion direction. CD: the normalized protrusion direction vectors are added as a combined vector and its magnitude is correlated with that of the cluster speed as CD. This parameter indicates if the combined protrusion direction associates with the cluster speed magnitude. Again, CD is higher for running than rotating. CS: the correlation between protrusion mass and cluster speed, indicating whether the size of a protrusion is associated with cluster speed.

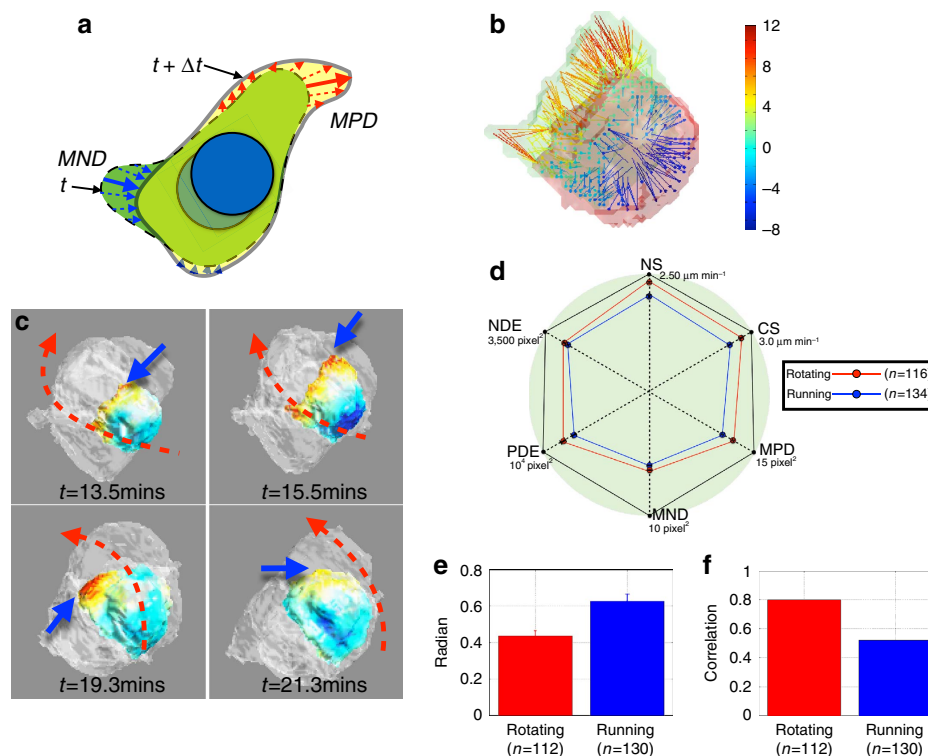
The association between cluster internal protrusions and cluster rotation is also validated. The rotating movement of the cluster takes into consideration all cell movements (represented by the nuclei center displacement vector) indicated by the blue arrow in Supplementary Fig. 12. The direction of a given cell to the cluster center is indicated by the red dashed line. The cross product of these two vectors gives a perpendicular vector and the combination of all those perpendicular vectors of all cells produces a combined cluster rotating vector indicated by the red arrow in Supplementary Fig. 12. Similarly, the cell positive deformation represents the cluster internal protrusion, illustrated by the green arrows in Supplementary Fig. 12. Applying the same computation to those vectors, we can get a combined protrusion rotating vector, represented by the magnet arrow in Supplementary Fig. 12. First, we checked the angle between combined cluster rotating vector and the combined protrusion rotating vector. The results of running and rotating mode are shown in Fig. 6e. We can see that the angle between these two vectors is  $<30$  degrees in rotating mode but much greater in running. We also validate that the rotating speed, represented by the magnitude of combined cluster rotating vector, is correlated with the protrusion amount, represented by the magnitude of combined protrusion rotating vector, in Fig. 6f. The correlation between these two parameters is about 0.80, compares to only 0.51 in running. Thus the cluster rotating behaviour is highly associated with the well-coordinated cluster internal protrusion. At times the cluster internal protrusion is

not coordinated, as shown in Supplementary Movie 9 and such behaviour is associated with neighbour exchange in the cluster.

## Discussion

The *Drosophila* border cell cluster is an excellent *in vivo* system to study collective cell migration. Quantitative analysis of the collectively migrating cells in full 3D is necessary to better understand the complex behaviour of individual cell and the cluster as a whole, but an appropriate computational solution has been lacking. Here we present a computational method including image pre-processing, segmentation, tracking and quantitative information extraction and data interoperability.

Although our computational solution is optimized, segmentation is somewhat limited by the imaging techniques. Although we may image the cluster with reasonable spacial and temporal resolution, we are still limited by current microscopy techniques. Typical cell membrane width is about  $0.01\ \mu\text{m}$  and such fine structure is not resolvable with the current imaging setup. To balance the spatial and temporal resolution, while reducing potential photo-toxicity, we can only achieve limited spatial resolution. A typical border cell has a volume of  $\sim 1,000\text{--}1,500\ \mu\text{m}^3$  ( $\sim 20,000\text{--}30,000$  voxels). This gives us an estimated 3D object with a radius of only 17–20 voxels. Although we enhance the weak membrane signal, segmentation errors may still happen. As the watershed is a method using the gradient of the signal. Both noise and other factors may affect the real gradient of the signal. For example, when two membranes



**Figure 6 | Cluster internal cell extensions are associated with cluster rotation and cells are more active during the rotating modes.** (a) Illustration of morphodynamics parameter calculations: the Maximum Positive Deformation (MPD), Maximum Negative Deformation (MND), Positive Deformation Energy (PDE) and Negative Deformation Energy (NDE). The solid red arrow is the MPD and the blue solid arrow is the MND. The PDE is the summation of all red arrows and the NDE is the summation of all blue arrows. (b) Morphodynamics of a border cell in 3D. Hot color indicates positive deformation and cold color indicates negative deformation. (c) Cluster internal cell extensions are associated with cluster rotation. Blue bold arrows indicate the cluster internal cell extensions of a border cell and the red dotted arrows indicate the rotating direction of the cluster. (d) Six measured parameters, including Nucleus Speed (NS), Cell Speed (CS) and those defined in a reveal that the cells are more active during the rotating mode than the running mode. Nucleus and cell speed are higher, and the cell front edges are more active, generating larger cluster internal cell extension during the rotation, while the retraction behaviour remains similar, indicated by the MND and NDE. (e) The angle difference between the cluster rotating vector (red arrow in Supplementary Fig. 12) and protrusion rotating vector (magnet arrow in Supplementary Fig. 12). The difference between these two vectors is often  $< 30$  degrees in rotating mode but much greater in running mode. Error bar represents s.e.m. (f) The correlation between the rotating speed and the positive protrusion amount. The correlation is about 0.80, compares to only 0.51 in running mode. This indicates cluster rotating behaviour is highly associated with the well-coordinated cluster internal protrusions.

of the same cell are very close, the nearby signal may affect the reading. Within the above imaging our method resolves nuclei, cell surfaces and their morphology as shown in Fig. 1.

The tracking of moving nuclei and cells is still a challenging step too, especially when cell speed is highly variable. Although the nuclei and cells can have significantly higher speed at some time points, the proposed tracking solution is able to provide us a metric which can be optimized using association algorithms and promises a global minimum. Based on our results, conventional parameters, such as volume and speed, can be extracted in the real 3D context on a cell-by-cell basis (Fig. 2). We observe that cell speed is generally higher than nucleus speed (Supplementary Fig. 4). The cell membrane is more active than the nucleus, which generally remains in a round or spherical shape. The slower speed of the nuclei is caused by the fact that the cell membrane is deforming, squeezing and moving forward, and hence displacing the cell's center of mass while the nucleus does not follow immediately. We observed two possible scenarios (Fig. 2d). First, the membrane is deforming and moving, while the nucleus does not move much. This will cause a difference between cell and nuclei speed. Second, both the membrane and the nucleus move together with a higher speed.

Previous qualitative observations of highly coordinated and chaotic movement of the cluster have been reported. The coordinated movement can be deconvolved into two distinct modes, running and rotating. These can be quantitatively described and classified using an experimental boundary. Discrete switching between these two modes occurs infrequently. The running and rotating modes in this *in vivo* study are very similar to a previous *in vitro* work<sup>20</sup> and similar collective behaviour has also been reported in schools of fish<sup>21</sup>. Generally, the motion of the cluster can be described as changing combinations. The running mode tends to dominate as the cluster initiates migration, whereas the rotating mode dominates as the cluster nears its destination<sup>2,8</sup>. In addition, the border cell cluster has some chaotic statuses between the rotating and the running modes, which remain difficult to distinguish from coordinated movement (Supplementary Fig. 8). Indeed, our quantitative analysis suggests a continuum of migration states between simple running and rotating classifications as in Fig. 3b,c. Our work also suggests that while simple parameters such as cluster speed may be useful to measure, additional insight can be gained by looking at the frequency of neighbour exchange using Angular Momentum and Group Polarity.

We demonstrate the possibilities enabled by our image analysis solution, and can make a few novel insights into the border cell migration. First of all, we show that polar cells rotate with the whole cluster during the rotating mode, while individual border cells are able to slide over the polar cell surface and exchange position with their neighbours (Fig. 4). Second, defining the surfaces of the individual nuclei/cells is crucial to understand the dynamics of cell-cell coupling and the degree of the coordination. Even without neighbour exchange, border cells can significantly change the contacting surface with other border cells during migration (Fig. 2f,g). Third, we are able to measure cluster external protrusions and quantify their correlation with cluster motion during running and rotating modes (Fig. 5). Finally, we identify another class of extensions, cluster internal cell extensions, which cannot easily be observed or quantified in 2D analysis at the whole cluster level (Fig. 6c). These extensions 'reach around' neighbouring border cells, and are associated with rotation of the cluster and neighbour exchange. Our 3D analysis at single cell resolution also allows us to show that the cells do not lose their mobility during the rotating mode. Study of cell membrane dynamics and the associated force is the key to understanding collective cell migration. Some solutions have been proposed for the computation of cell membrane deformation in 2D<sup>22,23</sup>. Based on our assumption that the sampling interval is small enough, we applied the linear mapping solution between two consecutive time points to estimate the deformation of a given cell in 3D. On the basis of our 3D morphological dynamics data, border cells are slightly more active in rotating than running. We hypothesize that during the rotating mode, the cluster may encounter less friction when rotating within the space it occupies than when running, in which the cluster must overcome the frictions at the interface between nurse cells to occupy new space. It is also possible that the rotating mode represents a highly dynamic stage of signal sensing between more progressive running modes. Further work will be required to test these hypotheses, for example by measuring tension in the different modes or using signaling pathway reporters to assess differences in signaling states. The mechanism by which the cluster transitions between the running and rotating modes remains unclear but the approaches described here provide a basis for its further investigation.

Our dedicated solution, the CCMTToolkit, provides a valuable new set of tools to automate the study of collective cell migration in 4D. It is freely available at (<https://sites.google.com/site/ccmtoolkit/>) with relevant technical documents, developer's API and testing data set. While our initial work has focused on wild type clusters, using these extra parameters, it will be possible to characterize different mutations and RNAi or overexpression phenotypes in more detail. For example, mutations which reduce border cell-border cell adhesion may display larger changes in neighbour interfaces, and may display higher rates of neighbour exchange. Taking an automated, computational approach will allow more objective interpretation of mutant behaviours and allow us to cluster mutant types based on behavioural similarities. Application of CCMTToolkit to wild type and mutant border cell clusters has the potential to offer novel insights into the mechanisms of collective cell migration. In general, this solution is applicable to other cell migration and tracking problems subject to optimization of imaging protocols.

## Methods

**Fly genetics and imaging approach.** UAS-*PH-PLCdelta*-GFP, UAS-*histone-RFP*, *Upd-Gal4* and *Slbo-Gal4* are described in Flybase [Flybase.org]. *UMAT-Lyn-tdTomato* was constructed by inserting the membrane-targeted *Lyn-tdTomato* fusion (gift from Darren Gilmour) into *pUMAT* (containing the maternal alpha tubulin promoter, gift from Daniel St. Johnston) and transgenic flies were made

by Bestgene. *Slbo-Gal4* drives UAS transgene expression in outer migratory border cells, and *Upd-Gal4* in polar cells. Together they allow expression of the nuclear marker (*histone-RFP*) and a cell membrane marker (*PH-GFP*) in all cluster cells. *UMAT-Lyn-tdTomato* marks nurse cell and oocyte membranes.

For two-marker videos: *UpdGal4/+; UAS-PH-GFP,slboGal4/UAS-his-RFP*. For three-marker videos: *UpdGal4/+; UMAT-Lyn-tdTom/+; UAS-PH-GFP,slboGal4/UAS-his-RFP*. For videos with single (outer) border cells marked: *hsFLP/+; 2x(tubp-GAL80), FRT80,slboGal4/UAS-PH-GFP,UAS-his-RFP,FRT80*. For videos of marked polar cells: *UpdGal4/+; UAS-PH-GFP,UAS-his-RFP/+*.

Dissection and live imaging of egg chambers was performed in imaging medium (Schneider's medium supplemented with 2.5% FCS, 5  $\mu\text{g ml}^{-1}$  insulin, 2  $\text{mg ml}^{-1}$  trehalose, 5  $\mu\text{M}$  methoprene, 1  $\mu\text{g ml}^{-1}$  20-hydroxyecdysone, 50  $\text{ng ml}^{-1}$  adenosine deaminase and 9  $\mu\text{M}$  FM 4-64) as described in ref. 24. Imaging was performed using an inverted confocal microscope (Leica SP5) For three-marker videos, four channels were acquired simultaneously: GFP (488 nm laser, collect emission 505–550 nm), Red 1 (561 nm laser, collect emission 570–620 nm), Red 2 (561 nm Laser, collect emission 620–790 nm) and transmission image (DIC). *tdTomato* and *dsRed* signals were unmixed using a linear unmixing approach. This was implemented using ImageJ and the Spectral Unmixing plugin written by Joachim Walther. The egg chamber drifting was corrected using the method in ref. 25. For two-marker videos, red 1 and red 2 channels were combined for RFP signal. The movies were recorded by using an inverted confocal microscope (SP5; Leica) with a 63 $\times$ , 1.2 NA Plan Apochromat water immersion objective with 2.5 zoom factor. The resolution is 296 $\times$ 150. Sampling intervals were 1–1.5 min and imaging continued until the border cells touch the edge of the imaging volume or a maximum of 2 h.

**Image reconstruction and pre-processing.** Confocal microscopy is able to acquire 3D time-lapse videos, however the image stacks may have some quality issues. The challenges include the following aspects: (a). Acquired image stacks are of low signal-noise-ratio. This is because moderate laser power is selected to minimize the photo-bleaching; (b). Despite moderate laser power, photo-bleaching still occurs, especially for the nuclei channel (Supplementary Fig. 1); (c). Anisotropic voxel size of different directions caused by: (i) The digital motor step in the *z* axis is generally larger than the *x-y* resolution; (ii) point spread function (PSF) has less optical resolution in the *z* axis.

Proper pre-processing is needed to address those quality issues. First of all, voxel size of acquired image stacks is different in different directions and we need to reconstruct them. The voxels size is normalized using linear interpolation and the scales of voxels are then isotropic post-reconstructed, that is, each voxel is 0.33 $\times$ 0.33 $\times$ 0.33  $\mu\text{m}$ . This step compensates for the bigger motor step in *z* direction, but cannot solve the issue of optical PSF. Based on the assumption that the acquired image contains a constant amount of fluorescence, the photo bleaching of all channels are compensated. The corrections of fluorescent energy for different channels are shown in Supplementary Fig. 1. This step potentially also amplified the noise at the latter stage of a movie. To reduce the noise, image stacks are smoothed by Gaussian kernel. Non-uniform background removal is also applied.

In our study, there are two types of cells within the migrating cluster: (i) A pair of polar cells that are in the center of the cluster; (ii) A fixed number of border cells surrounding the polar cells.

The only manual work in our analysis is to annotate the center of these two types of cells in the first image stack. For this we designed an annotation package named 'CCM Annotator' for this step. The remaining steps are fully automatic, including nuclei segmentation based on GMM, nuclei/cell tracking using a global cost function and cell segmentation based membrane signal.

**Nuclei segmentation based on GMM.** While previous efforts have been made on cell/nuclei segmentation and tracking it remains a challenging task. Existing segmentation approaches for nuclei segmentation, include simple thresholding<sup>26</sup>, watershed approach<sup>27,28</sup>, iterative voting methods<sup>29</sup>, level set approach based on gradient flow<sup>30</sup> and flexible contour model<sup>31</sup>. The watershed approach has been widely applied for cell segmentation. Different solutions have been proposed to overcome over-segmentation, for example, rule-based merging<sup>32</sup> and marker-controlled Voronoi diagram<sup>33</sup>, a level set formulation for Watershed segmentation<sup>34</sup>, preserving topology by simple point concept<sup>35</sup>. Some recent approaches are reported, such as multiple level set function for individual cells<sup>36</sup> and topological dependence<sup>37–39</sup>. Tracking is particularly challenging. Two basic point-based tracking techniques are centroid tracker<sup>40</sup> and correlation tracker<sup>41,42</sup>. Other tracking approaches include snake model with gradient vector flow<sup>43</sup>, snake model with shape and size constraint<sup>44,45</sup>, active contour approach for both segmentation and tracking<sup>46</sup>, K-mean clustering and matching<sup>47</sup>, and Gaussian Mixture Model<sup>48,49</sup>. Other relevant works on 3D time-lapse analysis include adaptive recursive analysis based on iterative thresholding<sup>50</sup>, blob detection based on Laplacian of Gaussian<sup>51</sup>, and a constrained active contours approach<sup>52</sup>. As we mentioned, application of those algorithms to the border cell system is not trivial and we are in need of a complete and automatic solution.

The *a priori* knowledge on the fixed cell number is a major challenge and we do not tolerate any segmentation or tracking error in a movie of typical  $\sim 100$  time points. Furthermore, we need to identify the location and boundaries of the nuclei

and cells. Since the nuclei are closely packed with each other, a Laplacian kernel is applied to enhance the boundaries between them.

From the annotation at the first time point, we know the number of nuclei and their identities, that is, polar or border cell. We denote a fixed number of nuclei using  $\omega_1, \omega_2, \dots, \omega_c$ , where  $c$  is the total number of nuclei. A threshold and some morphological operations are applied to obtain the foreground of the nuclei, for example, a mask covering all nuclei. During this process, we conserved the total mass of all nuclei.

For each voxel belonging to the foreground, we may form a vector, which is given by:

$$\vec{v}_i = f_n(x_i, y_i, z_i)(x_i, y_i, z_i) = w_i(x_i, y_i, z_i) \quad (1)$$

Where  $(x_i, y_i, z_i)$  are the coordinates of the given voxel and  $f_n(x_i, y_i, z_i)$  is the brightness at the voxel in the nuclei channel, denoted by  $w_i$ .  $w_i$  is just a weighting parameter for this coordinate. In total we have  $n$  such samples giving  $\vec{V} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ , where  $n$  is the total number of voxels in the mask. The segmentation of nuclei now becomes a task to assign a unique label to each sample from  $\omega_1, \omega_2, \dots, \omega_c$ .

We assume that the samples are drawn independently and identically from the distribution of  $p(\vec{v}|\omega_j)$ , which has a known parametric form that is uniquely determined by a parameter vector  $\vec{\theta}_j$ . Explicitly, we write  $p(\vec{v}|\omega_j)$  as  $p(\vec{v}|\omega_j, \vec{\theta}_j)$ . Here independent means the samples from different state  $\omega_j$  give no information about  $\theta_i$  if  $i \neq j$ . Here our assumption is that the parameters for different states are functionally independent. In our study, we have  $p(\vec{v}|\omega_j, \vec{\theta}_j) \sim N(\vec{\mu}_j, \Sigma_j)$  and  $\vec{\theta}_j$  includes both  $\vec{\mu}_j$  and  $\Sigma_j$ . Thus our observation  $\vec{V} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$  is drawn from a Gaussian Mixture Model (GMM), given by:

$$p(\vec{v}|\vec{\theta}) = \sum_{j=1}^c p(\vec{v}|\omega_j, \vec{\theta}_j)P(\omega_j) \quad (2)$$

$\vec{\theta} = (\vec{\theta}_1, \vec{\theta}_2, \dots, \vec{\theta}_c)$  is an unknown stationary parameter vector that we want to estimate from observations. In principle, if  $p(\vec{v}|\vec{\theta})$  is identifiable,  $\vec{\theta}$  can be recovered from empirical observations. Here identifiable indicates there exist a  $\vec{v}$  such that  $p(\vec{v}|\vec{\theta}) \neq p(\vec{v}|\vec{\theta}')$  if  $\vec{\theta} \neq \vec{\theta}'$ . The mixture of norm density is usually identifiable. The likelihood of observed samples is a joint density:

$$p(\vec{V}|\vec{\theta}) = \prod_{k=1}^n p(\vec{v}_k|\vec{\theta}) \quad (3)$$

The maximum likelihood (ML) estimation of  $\vec{\theta}$  is the value of  $\vec{\theta}$  that maximizes  $p(\vec{V}|\vec{\theta})$  in Equation (3). As we already know the number of cells, we only need to fit our data with the given model.

As we assumed  $p(\vec{v}|\omega_j, \vec{\theta}_j) \sim N(\vec{\mu}_j, \Sigma_j)$ , thus  $p(\vec{V}|\vec{\theta})$  is differentiable with respect to  $\vec{\theta}$ . We can actually derive the analytical necessary condition for  $\vec{\theta}$  since we know the gradient respect to  $\vec{\theta}$  must vanish at  $\vec{\theta}$ . However, when both  $\vec{\mu}_j$  and  $\Sigma_j$  are unknown in our study, the ML solution is singular. Although the number of the nuclei is fixed, we can let the variance approach zero thus make the likelihood arbitrarily large. Even so, a useful solution can still be obtained when placing meaningful constraints on the covariance matrix to ensure positive definiteness of the obtained covariance matrix. Since we are sure that the values at the diagonal entries of each covariance matrix should not be smaller than some value, a tiny positive real can be added to the diagonal entries as a regularizer to enforce this prior constraint. Or we can restrict ourselves to work with only diagonal covariance matrices. In practice, the classical expectation-maximization (EM) algorithm is considered to obtain the local maximal solution of the likelihood function in Equation (3). Once the ML solution for  $\vec{\theta}$  is obtained, in our context we have ready each of the components  $\vec{\mu}_j$  as the weighted geometric center of each nuclei, together with the corresponding shapes described by  $\Sigma_j$ .

**Nuclei/cell tracking with global cost function optimization.** The tracking of nuclei is a problem of association or matching at different time point. It can be solved using Markov Chain or traditional point match algorithms. In this paper, we propose an appearance based tracking model using a global cost function optimization.

Let's say the nuclei segments at time  $t$  and  $t+1$  are denoted by  $\omega_i^t$  and  $\omega_i^{t+1}$  respectively, where  $i = 1, 2, \dots, c$ . We can define a matrix of tracking confidence and each element in this matrix is given by:

$$t_{ij} = \frac{\|\omega_i^t \cap \omega_j^{t+1}\|}{|\mu_i^t - \mu_j^{t+1}| + \varepsilon} \quad (4)$$

$\|\cdot\|$  in the numerator means the norm of a set of voxels, which give us the overlapped volume of two segments at two different time points.  $|\cdot|$  in the denominator means the Euclidean distance between the nuclei centers at two time points. Intuitively, Equation (4) tells us if the distance between two segments is smaller, they are more likely associated and it also favours the segments, which have greater overlapping volume.  $\varepsilon$  is a regulator to increase the stability of the matrix.

In building such a matrix of tracking confidence, the nuclei tracking becomes an assignment problem. Given two sets  $\Omega^t = \{\omega_i^t\}$  and  $\Omega^{t+1} = \{\omega_i^{t+1}\}$  of equal size, together with the tracking confidence  $T: \Omega^t \times \Omega^{t+1} \rightarrow R$ , which elements are

defined in Equation (4). We hope to find a bijection (also known as one-to-one matching.)  $f: \Omega^t \rightarrow \Omega^{t+1}$  such that cost function  $\sum_{\omega_i^t \in \Omega^t} T(\omega_i^t, f(\omega_i^t))$  is maximized, which gives us:

$$\arg \max_f \left( \sum_{\omega_i^t \in \Omega^t} T(\omega_i^t, f(\omega_i^t)) \right) \quad (5)$$

Equation (5) means that of all possible matches, we hope to find a perfect match, which maximizes the summation of our tracking confidence at  $t+1$ . This problem can be efficiently solved using combinatorial optimization algorithms, such as the classical Hungarian algorithm. This guarantees an optimal solution for the cluster of cells, instead of just using the local information for nuclei tracking. As we can see in Supplementary Movie 1, even when some cells move very fast at certain times, we can still track them successfully using our approach.

**Marker controlled cell segmentation using membrane signal.** We know that the cluster and/or individual cells do not suddenly gain or lose mass during the migration process. Constraining the volume of each individual cell is theoretically feasible, but given the limitations of image acquisition such a solution may not bring us more benefit and is much more computationally expensive. In our work, we used a greedy algorithm to determine an optimal thresholding value which provides a mask for the cluster with a volume determined by the first time point.

The membrane signal is relatively weak and lacking in specificity so we used a rotational kernel to enhance the membrane signal based on the infinity norm. Since the nuclei are already segmented and tracked, we can then use them as seeds to apply a seeds-controlled watershed algorithm to segment the cells based on the enhanced membrane signal.

**The coordination of nuclei/cell movement.** After nuclei tracking, we can quantify nuclei movement. Let's say nucleus  $\omega_{i,t}^n$  at time point  $t$  matched to  $\omega_{j,t+1}^n$  at time point  $t+1$ . Their geometric centers are represented by  $(x_i^t, y_i^t, z_i^t)$  and  $(x_j^{t+1}, y_j^{t+1}, z_j^{t+1})$ , then we may calculate a vector to represent the movement of each nucleus:

$$\vec{V}_{i,j}^{t+1} = (x_i^t - x_j^{t+1}, y_i^t - y_j^{t+1}, z_i^t - z_j^{t+1}) \quad (6)$$

The displacement, that is, magnitude, between two time points:

$$D_{i,j}^{t+1} = |(x_i^t - x_j^{t+1}, y_i^t - y_j^{t+1}, z_i^t - z_j^{t+1})| \quad (7)$$

Where  $|\cdot|$  means norm. Note that  $\vec{V}_{i,j}^{t+1}$  is a vector and its magnitude is given by

$$v_{i,j}^m = \frac{D_{i,j}^{t+1}}{\Delta t} \quad (8)$$

The direction of  $\vec{V}_{i,j}^{t+1}$  can be represented by a normalized vector:

$$\vec{v}_{i,j}^d = \frac{(x_i^t - x_j^{t+1}, y_i^t - y_j^{t+1}, z_i^t - z_j^{t+1})}{D_{i,j}^{t+1}} \quad (9)$$

Note that  $v_{i,j}^m$  is scalar, but  $\vec{v}_{i,j}^d$  is a vector with norm equal to 1. At each time point, we will have a number of tracking identities. The concept of velocity is a vector which is represented by its magnitude and direction. For simplicity, let's use  $v_k^m(t)$  and  $\vec{v}_k^d(t)$ , where  $k = 1, 2, 3, \dots$ , to represent its magnitude and direction respectively.

We visualized the nuclei velocity direction with the original image stacks. We observed clearly two different migration statuses, called running mode and rotating mode, illustrated in Fig. 3b,c. Furthermore, we calculated two features, named Group Polarization and Angular Momentum, for a quantitative understanding of these two different migrating statuses.

$\vec{v}_{k,t+1}^d$  is a vector only representing the direction of each nucleus' movement and contains no magnitude information. We interrogate the question of whether the nuclei move in the same direction at each time point or are moving in different directions. To answer this question, we may calculate the summation of  $\vec{v}_{k,t+1}^d$  at each time point and denoted it as Group Polarization:

$$P(t) = \frac{\left| \sum_k \vec{v}_{k,t}^d \right|}{K} \quad (10)$$

Where  $K$  is total number of  $\vec{v}_{k,t+1}^d$  and  $P(t)$  is a scalar between  $[0, 1]$ . If the cluster is in the running phase, then  $P(t)$  is close to 1.0 as shown in Fig. 3a. The value of  $P(t)$  will be much lower when the cluster is under rotation phase. The calculation of  $P(t)$  is represented by the red curve in Fig. 3a.

To identify the rotation movement of the cluster, we further define the angular momentum of the cluster. For each nucleus, its center position is denoted as  $c_i(t)$ . The cluster center is then given by

$$C(t) = \frac{1}{K} \sum_i c_i(t) \quad (11)$$

We obtain the vector of the nuclei to the cluster center as:

$$\vec{r}_i(t) = \frac{c_i(t) - C(t)}{|c_i(t) - C(t)|} \quad (12)$$



Note that  $\vec{r}_i(t)$  is a normalized vector. The angular momentum is given by:

$$M(t) = \frac{1}{K} \left| \sum_{i=1}^K \vec{r}_i(t) \times \vec{v}_i^d(t) \right| \quad (13)$$

Again  $K$  is total number of  $\vec{r}_i(t)$  and  $M(t)$  is a scalar between  $[0, 1]$ . If the cluster is in the rotation phase,  $M(t)$  will be close to 1.0 as shown in Fig. 3a. The value of  $M(t)$  will be much lower when the cluster is under running phase. The calculation of  $M(t)$  is represented by the blue curve in Fig. 3a.

**Neighbour exchange of border cells during migration.** Another interesting behaviour we aimed to quantify is neighbour exchange. An example of which is shown in Fig. 4a. The nucleus/cell center position, denoted by  $c_i(t)$ , can provide such information. We first calculate the inter-cell distance matrix at a given time  $t$  as  $D_t = (d_{ij})_{n \times n}$ . Here  $d_{ij} = |c_i(t) - c_j(t)|$ . Let  $\tau \in \{1, 2, 3, \dots\}$  be the temporal stepsize, and define the temporal difference of step-size  $\tau$  at time point as:

$$\delta_t^{(\tau)} := \max(|D_{t+\tau} - D_t|) \quad (14)$$

With  $|\bullet|$  denoting element-wise absolute values and  $\max(\cdot)$  being the max operation over the matrix element.

Empirically the measurement  $\delta_t^{(\tau)}$  are plotting over time  $t$  and the temporal scale spaces  $\tau$ . Denote the cut-off thresholding  $\rho := 60$ . Clearly most of the neighbourhood exchange cases have very high values ( $\geq \rho$ ), while most of the rotation and sliding cases have relatively low values ( $< \rho$ ). This suggests  $\delta_t^{(\tau)}$  is a good parameter to detect the neighbourhood exchange during collective cell migration. It is worth noting that the threshold  $\rho := 60$  is about the twice the average nucleus diameter (30 voxels) in our context. Meanwhile, the temporal scale space offers new insights into how the neighbourhood exchanges could be organized over time.

**Cluster external protrusion detection and analysis.** The protrusions formed by the border cells play an important role in the migratory process. There are some specific biomarkers for different protrusions but these are difficult to implement in the *in vivo* experiments. Here we use a morphological operation to detect cell extensions. Typically, a cluster external protrusion is a structure extending along the normal direction of the cluster surface and its thickness is smaller than the diameter of the cluster. Though cluster external protrusions are cluster level structures, they are still generated by individual border cells. We calculated the averaged volume of the border cell and use this volume to fit it as a sphere. The radius of this sphere is then used for a morphological operation, called opening, which is given as follows. If the binary image of the cluster is given by  $B$ , and the ball structure element is given by  $S$ , then the detected cluster external protrusion is:

$$P = B - (B \circ S) = B - ((B \bullet S) \oplus S) \quad (15)$$

Where  $\circ$  is the opening operation, which is erosion, that is,  $\bullet$ , followed by dilation  $\oplus$  of the given element ball structure of radius  $r$ . Some parameters of the detected cluster external protrusions are measured. Cluster external protrusion size is calculated by  $|P|/\Delta t$ , where  $|P|$  is taking the norm of the protrusion, that is, its volume and  $\Delta t$  is the sampling interval between the two image stack. The normalization according to the sampling interval is necessary, as if we see a bigger protrusion size, we are not sure that if it is caused by more active front cell behaviour or it has longer time to generate bigger protrusion. The cluster external protrusion's direction is measured according to difference between the cluster mass center and the protrusion mass center. During the tracking, the speed of cluster mass center is calculated as follows. After the opening operation, that is,  $(B \circ S)$ , we actually take away the elongated protrusions, and the rest is a more spheroid ball, we called cluster body. Let  $\mathbf{m}(t) = (x, y, z)$  is the position of cluster body mass center,  $\mathbf{m}(t+1) = (x', y', z')$  is for  $t+1$ . The  $\mathbf{d}(t)$  is the displacement between  $\mathbf{m}(t)$  and  $\mathbf{m}(t+1)$ . Then cluster speed is given by  $\mathbf{v}(t) = \mathbf{d}(t)/\Delta t$ . The direction of this vector defines the cluster direction. The detected cluster external protrusions of a given movie is shown in Supplementary Movie 10.

The correlation is a standard parameter to measure how strongly two variables are correlated. Let's say the two variables are  $X(t)$ , that is, combined directions and  $Y(t)$ , i.e. cluster speed magnitude. The correlation matrix is calculated in this way:  $\Sigma = (E(X - \mu_x)(Y - \mu_y))/(\sigma_x \sigma_y)$ . Where  $\mu_x$  and  $\mu_y$  is the mean of  $X$  and  $Y$ ;  $\sigma_x$  and  $\sigma_y$  is the standard deviation of  $X$  and  $Y$ . Note that  $\Sigma$  is a  $2 \times 2$  symmetric matrix. The diagonal elements are the correlation coefficient between  $X$  and  $Y$ . If this correlation is equal to 1, it means the two variables are positive linearly correlated; if it is zero, it means the two variables have no relationship to each other. Supplementary Movies 11–12 illustrate the deformation of the border cells in running and rotating modes.

**Computational 3D morphological dynamics.** In our experiments, the high temporal resolution of the raw data as well as our accurate segmentation provided us with a basis to look at the dynamics of each migrating cell surface over time in 3D, that is the morphodynamics. The movement of the cell is a continuous process and the dynamics of its surface contains a lot of information. Some methods are developed to explore the cell membrane deformation in 2D are reported<sup>22,23</sup>. In our 3D study, it is to build an association of two cell surfaces between two consecutive time points. Once we achieve this, we can then estimate the deformation along the cell surface at two different time points. The retraction and

forward extension of the cell can thus be quantitatively described. For a migrating and deforming cell, its surface can be described as a function, that is,  $S(x, y, z, t)$ . In Fig. 6b, we demonstrate our 3D computation in 2D. As we have excellent temporal resolution, we assume that any deformation along the norm direction of a cell surface is small enough that we can treat it as linear. The deformation between any two time points is given by:

$$\Delta S = \frac{\partial S}{\partial t} \quad (16)$$

On the basis of our linear assumption, Equation (16) is actually to find a linear minimum distance mapping for the points on the surfaces at two different time points. The arrows in Fig. 6b, show the matching points. We also need to define whether the deformation is positive (extending) or negative (retracting). Positive deformation means a surface moving into a space which was not previously occupied by the cell, shown by the red arrows in Fig. 6b; while negative deformation means a surface retracting into a space which was previously internal to the cell, shown as the blue arrow in Fig. 6b. The actual deformation calculation of a given cell is shown in Fig. 6c. In our work, we take the positive deformation of the cell surface as the cluster internal protrusion.

To quantify the cell mobility, we defined the following 6 parameters. Nuclei speed: represented as 'NS' axis Fig. 6d; cell speed: represented as 'CS' axis in Fig. 6d; maximum positive deformation: The maximum deformation vector size in the positive deformation vectors, shown as the solid red arrow in Fig. 6a represented as 'MPD' axis in Fig. 6d; maximum negative deformation: The maximum deformation vector size in the negative deformation vectors, shown as the solid blue arrow in Fig. 6a represented as 'MND' axis in Fig. 6d; positive deformation energy: The integration of all positive deformation, that is, the summation of all red arrows as shown in Fig. 6a, represented as 'PDE' axis in Fig. 6d; negative deformation energy: The integration of all negative deformation, that is, the summation of all blue arrows as shown in Fig. 6a, represented as 'NDE' axis in Fig. 6d.

**Data availability.** The data sets generated during and/or analysed during this study are available in the website of CCMTolKit (<https://sites.google.com/site/ccmtolkit/>). Both source code and a few example movies are provided. Other further information and details are available from the corresponding author on reasonable request.

## References

- Friedl, P. & Gilmour, D. Collective cell migration in morphogenesis, regeneration and cancer. *Nat. Rev. Mol. Cell. Biol.* **10**, 445–457 (2009).
- Rorth, P. Collective Cell Migration. *Annu. Rev. Cell Dev. Biol.* **25**, 407–429 (2009).
- Ridley, A. J. *et al.* Cell migration: integrating signals from front to back. *Science* **302**, 1704–1709 (2003).
- Friedl, P., Sahai, E., Weiss, S. & Yamada, K. M. New dimensions in cell migration. *Nat. Rev. Mol. Cell. Biol.* **13**, 743–747 (2012).
- Petrie, R. J. & Yamada, K. M. At the leading edge of three-dimensional cell migration. *J. Cell Sci.* **125**, 5917–5926 (2013).
- Montell, D. J. Border-cell migration: the race is on. *Nat. Rev. Mol. Cell. Biol.* **4**, 13–24 (2003).
- Bianco, A. *et al.* Two distinct modes of guidance signalling during collective migration of border cells. *Nature* **448**, 362–365 (2007).
- Inaki, M., Vishnu, S., Cliffe, A. & Rorth, P. Effective guidance of collective migration based on differences in cell states. *Proc Natl Acad. Sci. USA* **109**, 2027–2032 (2012).
- Poukkula, M., Cliffe, A., Chagnede, R. & Rorth, P. Cell behaviors regulated by guidance cues in collective migration of border cells. *J. Cell Biol.* **192**, 513–524 (2011).
- Prasad, M. & Montell, D. J. Cellular and molecular mechanisms of border cell migration analyzed using time-lapse live-cell imaging. *Dev. Cell* **12**, 997–1005 (2007).
- Niewiadomska, P., Godt, D. & Tepass, U. DE-cadherin is required for intercellular motility during drosophila oogenesis. *J. Cell Biol.* **144**, 533–547 (1999).
- Wang, X., He, L., Wu, Y. L., Hahn, K. M. & Montell, D. J. Light-mediated activation reveals a key role for Rac in collective guidance of cell movement in vivo. *Nat. Cell Biol.* **12**, 591–597 (2010).
- Cai, D. *et al.* Mechanical feedback through E-cadherin promotes direction sensing during collective cell migration. *Cell* **157**, 1146–1159 (2014).
- McMahon, A., Supatto, W., Fraser, S. E. & Stathopoulos, A. Dynamic analyses of drosophila gastrulation provide insights into collective cell migration. *Science* **322**, 1546–1550 (2008).
- Xiong, F. *et al.* Interplay of cell shape and division orientation promotes robust morphogenesis of developing epithelia. *Cell* **159**, 415–427 (2014).
- Gelbart, M. A. *et al.* Volume conservation principle involved in cell lengthening and nucleus movement during tissue morphogenesis. *Proc. Natl Acad. Sci. USA* **109**, 19298–19303 (2012).
- Yoshida, S. *et al.* Genetic control of plant development by overriding a geometric division rule. *Dev. Cell* **29**, 75–87 (2014).

18. Barry, D. J., Durkin, C. H., Abella, J. V. & Way, M. Open source software for quantification of cell migration, protrusions, and fluorescence intensities. *J. Cell Biol.* **209**, 163–180 (2015).
19. Mosaliganti, K. R., Noche, R. R., Xiong, F., Swinburne, I. A. & Megason, S. G. ACME: automated cell morphology extractor for comprehensive reconstruction of cell membranes. *PLoS Comput. Biol.* **8**, e1002780 (2012).
20. Malet-Engra, G. *et al.* Collective cell motility promotes chemotactic prowess and resistance to chemorepulsion. *Curr. Biol.* **25**, 242–250 (2015).
21. Tunström, K. *et al.* Collective States, multistability and transitional behavior in schooling fish. *PLoS Comput. Biol.* **9**, e1002915 (2013).
22. Tyson, R. A., Epstein, D. B. A., Anderson, K. I. & Bretschneider, T. High resolution tracking of cell membrane dynamics in moving cells: an electrifying approach. *Math. Model. Nat. Phenom.* **5**, 34–55 (2010).
23. Ji, L., Lim, J. & Danuser, G. Fluctuations of intracellular forces during cell protrusion. *Nat. Cell Biol.* **10**, 1393–1400 (2008).
24. Poukkula, M., Cliffe, A., Changede, R. & Rørth, P. Cell behaviors regulated by guidance cues in collective migration of border cells. *J. Cell Biol.* **192**, 513–524 (2011).
25. Cliffe, A., Poukkula, M. & Rørth, P. Culturing *Drosophila* egg chambers and imaging border cell migration. *Nat. Protoc. Exchange* doi:10.1038/nprot.2007.289 (2007).
26. Lerner, B., Clocksin, W. F., Dhanjal, S., Hultén, M. A. & Bishop, C. M. Automatic signal classification in fluorescence in situ hybridization images. *Cytometry* **43**, 87–93 (2001).
27. Lockett, S. J. & Herman, B. Automatic detection of clustered, fluorescent-stained nuclei by digital image-based cytometry. *Cytometry* **17**, 1–12 (1994).
28. Malpica, N. *et al.* Applying watershed algorithms to the segmentation of clustered nuclei. *Cytometry* **28**, 289–297 (1997).
29. Parvin, B. *et al.* Iterative voting for inference of structural saliency and characterization of subcellular events. *IEEE Trans. Image Process.* **16**, 615–623 (2007).
30. Ortiz De Solorzano, C., Malladi, R., Lelièvre, S. A. & Lockett, S. J. Segmentation of nuclei and cells using membrane related protein markers. *J. Microsc.* **201**, 404–415 (2001).
31. Clocksin, W. F. in *Proceedings of 12th International Conference on Image Analysis and Processing* 682–687 (IEEE, 2003).
32. Wählby, C., Lindblad, J., Vondrus, M., Bengtsson, E. & Björkstén, L. Algorithms for cytoplasm segmentation of fluorescence labelled cells. *Analyt. Cell. Pathol.* **24**, 101–111 (2002).
33. Zhou, X., Liu, K. Y., Bradley, P., Perrimon, N. & Wong, S. T. C. Towards automated cellular image segmentation for rai genome-wide screening. *Med. Image Comput. Comput. Assist. Interv.* **8**, 885–892 (2005).
34. Tai, X.-C. *et al.* in *International Conference on Scale Space and Variational Methods in Computer Vision* (Springer Berlin Heidelberg, 2007).
35. Xiao, H., Chenyang, X. & Prince, J. L. in *Computer Vision and Pattern Recognition (CVPR). Proceedings of the 2001 IEEE Computer Society Conference* (IEEE, 2001).
36. Yan, P., Zhou, X., Shah, M. & Wong, S. T. C. Automatic segmentation of high-throughput RNAi fluorescent cellular images. *IEEE Trans. Inform. Technol. Biomed.* **12**, 109–117 (2008).
37. Yu, W., Lee, H. K., Hariharan, S., Bu, W. & Ahmed, S. Quantitative neurite outgrowth measurement based on image segmentation with topological dependence. *Cytometry A* **75**, 289–297 (2009).
38. Yu, W., Lee, H. K., Hariharan, S., Bu, W. & Ahmed, S. Evolving generalized Voronoi diagrams for accurate cellular image segmentation. *Cytometry A* **77A**, 379–386 (2010).
39. Yu, W. *et al.* In: *Advances in Visual Computing*. Vol. 5875 (eds Bebis G, *et al.*) pp 531–543 (Springer, 2009).
40. Ghosh, R. N. & Webb, W. W. Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules. *Biophys. J.* **66**, 1301–1318 (1994).
41. Acton, S. T., Wethmar, K. & Ley, K. Automatic tracking of rolling leukocytes in vivo. *Microvasc. Res.* **63**, 139–148 (2002).
42. Schütz, G. J., Schindler, H. & Schmidt, T. Single-molecule microscopy on model membranes reveals anomalous diffusion. *Biophys. J.* **73**, 1073–1080 (1997).
43. Seroussi, I., Veikherman, D., Ofer, N., Yehudai-Resheff, S. & Keren, K. Segmentation and tracking of live cells in phase-contrast images using directional gradient vector flow for snakes. *J. Microsc.* **247**, 137–146 (2012).
44. Ray, N., Acton, S. T. & Ley, K. Tracking leukocytes in vivo with shape and size constrained active contours. *IEEE Trans. Med. Imaging* **21**, 1222–1235 (2002).
45. Ray, N. & Acton, S. T. Motion gradient vector flow: an external force for tracking rolling leukocytes with shape and size constrained active contours. *IEEE Trans. Med. Imaging* **23**, 1466–1478 (2004).
46. Zimmer, C., Labruyere, E., Meas-Yedid, V., Guillen, N. & Olivo-Marin, J. C. Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: a tool for cell-based drug testing. *IEEE Trans. Med. Imaging* **21**, 1212–1221 (2002).
47. Xiaowei, C., Xiaobo, Z. & Wong, S. T. C. Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy. *IEEE Trans. Biomed. Eng.* **53**, 762–766 (2006).
48. Pham, T. D. & Tran, D. T. In: *2006 IEEE/SMC International Conference on System of Systems Engineering*, (IEEE, 2006).
49. Nogueira, P. A. & Teófilo, L. F. A multi-layered segmentation method for nucleus detection in highly clustered microscopy imaging: a practical application and validation using human U2OS cytoplasm–nucleus translocation images. *Artif. Intell. Rev.* **42**, 331–346 (2014).
50. Keller, P. J., Schmidt, A. D., Wittbrodt, J. & Stelzer, E. H. K. Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science* **322**, 1065–1069 (2008).
51. Keller, P. J. *et al.* Fast, high-contrast imaging of animal development with scanned light sheet-based structured-illumination microscopy. *Nat. Methods* **7**, 637–642 (2010).
52. Schmid, B. *et al.* High-speed panoramic light-sheet microscopy reveals global endodermal cell dynamics. *Nat. Commun.* **4**, 2207 (2013).

## Acknowledgements

We thank the funding support of Career Development Grant from JCO, A\*STAR, Singapore. We are also grateful for guidance from Pernille Rørth and her kind permission to use the data generated from her lab. We thank Giorgio Scita, Nicolas Plachta and Graham Wright for their invaluable feedback and thoughtful discussions on the manuscript.

## Author contributions

A.C., D.P.D., H.S. and I.L.K.H. prepared the *Drosophila* assays and performed the live imaging. A.C. designed and optimized the live imaging protocol. L.C. contributed the temporal topology change metric to detect the neighbour exchange. W.Y. designed the computation methods. K.H.O. contributed to the preparation of graphics, specially Figs 4–6 and the Supplementary Movies. W.Y. prepared the manuscript with the input from all other authors.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Cliffe, A. *et al.* Quantitative 3D analysis of complex single border cell behaviors in coordinated collective cell migration. *Nat. Commun.* **8**, 14905 doi: 10.1038/ncomms14905 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017